5Д

УДК 004.89

Р.Р. Даревич, Д.Г. Досин, В.В. Литвин, А.О. Мазур
Фізико-механічний інститут ім. Г.В. Карпенка НАН України, м. Львів, Україна
dosyn@ipm.lviv.ua, darevych@ipm.lviv.ua, vasyll@ukr.net, ania@ipm.lviv.ua

Віртуальна бібліотека – елемент інтелектуальної системи для координації наукових досліджень

Стаття присвячена створенню віртуальної бібліотеки в рамках побудови у Фізико-механічному інституті НАН України віртуального автоматизованого робочого місця наукового працівника. В основу електронного каталогу покладено стандарт метаданих Дублінське ядро (Dublin Core). Наповнена науковими публікаціями система створює передумови для здійснення автоматичного навчання онтології метапошукової системи у відповідній до потреб користувача галузі в рамках ВАРМ наукового працівника.

Вступ

На сучасному етапі побудови інформаційного суспільства серед тенденцій розвитку, зокрема, бібліотечної та видавничої галузей відокремлюються ті, які пов’язані з активним використанням електронного середовища та формуванням систем електронних інформаційних ресурсів.

Бібліотеки, видавництва, органи НТІ, наукові установи й навчальні заклади країни ведуть роботу зі створення електронних інформаційних ресурсів: формуються електронні каталоги, картотеки праць науковців, створюються архіви повнотекстових електронних документів. Ці інформаційні ресурси (ІР) потрібні широкому загалу науковців, але здебільшого вони існують лише в локальних мережах установ і не доступні через глобальну комп’ютерну мережу [1].

Від організації ІР та їх застосування залежить і прозорість, і своєчасність, і компетентність управлінських рішень. Також для підвищення ефективності наукових досліджень необхідно вести оперативний інформаційний пошук, облік ресурсів та забезпечити об’єктивний аналіз і оцінку результатів досліджень. Сучасний розвиток інформаційних технологій створює необхідні передумови впровадження програмно-технічних систем, здатних автоматизувати розв’язання цих задач [2]. Зокрема таким засобом є віртуальне автоматизоване робоче місце (ВАРМ) наукового працівника [3], що являє собою систему авторизованого доступу користувача-спеціаліста до корпоративної бази даних установи та відповідного його правам набору інформаційних служб через Інтранет і/або Інтернет-мережу для автоматизації виконання ним своїх професійних обов’язків. На сьогодні системи такого типу набувають широкого розповсюдження в повсякденній роботі спеціалістів різних галузей науки та промисловості [4]. Наприклад, автоматизоване робоче місце (АРМ) інженера-дослідника [5], навчальне АРМ перекладача [6], АРМ викладача з інтелектуальною підтримкою [7] та інші. Принципова відмінність ВАРМ від АРМ – можливість доступу користувача до системи з будь-якого комп’ютера, підключеного до мережі Інтернет, після авторизації, а також інформаційне забезпечення служб ВАРМ (забезпечення актуальності корпоративних баз даних та баз знань) уповноваженими спеціалізованими підрозділами.

У випадку розроблення подібної системи для наукової установи одним з ключових завдань постає забезпечення науковців актуальною спеціалізованою інформацією через її автоматичне поновлення, зокрема відповідальними службами наукової установи. Тому **метою роботи** є розробка програмних засобів інтеграції та координації досліджень – створення електронного каталогу та наповнення віртуальної бібліотеки електронними публікаціями в рамках віртуального автоматизованого робочого місця на базі мережевих та інформаційних технологій.

Структура ВАРМ наукового працівника

До складу ВАРМ, розробленого у ФМІ, входить ряд інформаційних служб (рис. 1).

Віртуальне автоматизоване робоче місце забезпечує користувача послугами файлового сервера, доступними через простий інтуїтивно зрозумілий веб-інтерфейс та авторизований доступ. Це − універсальна система, і у даному випадку вона може бути організована як віртуальна бібліотека, яка дозволяє зберігати електронні публікації у систематизованому згідно з УДК каталозі, доступна повністю через протокол HTTP та забезпечує три рівні доступу: «для всіх», «для зареєстрованих», для користувачів із заданим логіном та відповідним логіну паролем [8].

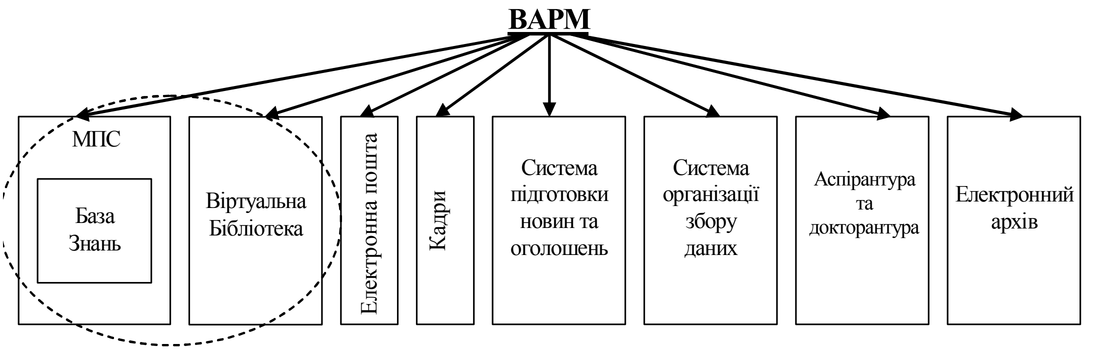


Рисунок 1 – Компоненти системи ВАРМ

Завдання віртуальної бібліотеки – надати «віртуальні полиці», іменовані за УДК, для зберігання та обміну науковими публікаціями та посиланнями на них. Як уже згадувалося в обґрунтуванні доцільності застосування ВАРМ у науковій сфері, наявність електронної віртуальної бібліотеки в науковій установі як відповідної послуги у складі ВАРМ дозволяє заощаджувати час та кошти завдяки тому, що один раз знайдений та придбаний документ (стаття, патент і т.п.) не буде замовлений установою вдруге, а буде надаватися на вимогу з локальної електронної бібліотеки. При цьому автоматично визначається популярність тих чи інших журналів та конкретних наукових праць, що може бути корисним при замовленні подальшої передплати, а також адаптації системи до потреб користувачів.

У рамках науково-дослідних робіт у ФМІ НАН України розробляється онтологія та наповнюється відповідна їй база знань експертної системи у галузі матеріалознавства, що забезпечить цей науковий напрям засобами збереження, пошуку, представлення, формалізації та стандартизації фахових знань. Віртуальна бібліотека інституту, яка входить до складу її Інтернет-порталу, наповнюється електронними публікаціями з цього наукового напряму.

Нижче розглянемо існуючі підходи та стандарти представлення даних у бібліотеці.

Аналіз підходів. Метадані та Дублінське ядро

Наявний досвід створення електронних бібліотек за кордоном та в Україні свідчить, що роботи з цієї проблематики сьогодні не мають належної системності й не забезпечують комплексного вирішення завдань правового, організаційного й технологічного характеру. Тим часом, це необхідно для розв'язання в цілому завдань побудови електронних бібліотек і їх мереж, тому принципово важливо вже на початковій стадії вирішення проблем створення в Україні електронних бібліотек застосувати системний підхід, що потребує їх аналізу як підсистеми певної більшої структури.

У процесі розробки електронної бібліотеки можна скористатися досвідом Національної бібліотеки України ім. Вернадського (НБУВ). Фахівці НБУВ провели аналіз роботи свого сайту з віддаленими абонентами, технологію дослідження динаміки відвідувань сайту та обробки помилкових запитів, впроваджену на Web-вузлі [9]. З моніторингу роботи сайту НБУВ випливає, що підвищеним попитом серед Internet-абонентів користуються реферативна інформація та повнотекстові матеріали, проте існує мало електронних версій книжок із фондів книгозбірні.

Також на сьогоднішній день правила бібліотечного опису (ГОСТ-7.1.84) не мають у своєму розпорядженні даних про цифрові ресурси. А оскільки бібліотечний каталог організує у єдину систему одиниці зберігання бібліотечних фондів, а бібліографічна картка є не що інше, як набір метаданих на книгу чи статтю з журналу, яка побудована за правилами бібліографічного опису, можна скористатися відомим стандартом метаданих Дублінське ядро (Dublin Core) [10]. Цей стандарт є форматом опису практично будь-яких інформаційних ресурсів, зокрема Інтернет – не складний за структурою, відносно легкий у застосуванні, інтернаціональний, міжнародно-визнаний.

Формат Dublin Core містить у собі 15 елементів для опису цифрового ресурсу:

Назва – Title,
Створювач (Автор) – Creator,
Тема (Предметна рубрика) – Subject,
Опис (Анотація) – Description,
Видавець – Publisher,
Співавтор – Contributor,
Дата – Date,
Формат – Format,
Тип – Type,
Ідентифікатор – Identifier,
Джерело – Source,
Мова – Language,
Відношення – Relation,
Охоплення (Покриття) – Coverage,
Авторські права – Rights.

У вересні 2001 р. набір метаданих Dublin Core був затверджений у США Американським Інститутом Національних Стандартів як стандарт Z39.85. В Австралії, Канаді, Данії, Фінляндії, Ірландії і Великобританії в цьому ж році формат Dublin Core рекомендований і прийнятий як державний стандарт для он-лайн ресурсів та е-комерції. З кінця 90-х почав широко використовуватися в Україні. Так, наприклад, наприкінці 1999 р. було прийняте рішення використовувати Dublin Core для цифрових ресурсів, що публікуються на сайті Кіровоградської ОУНБ [11]. Починаючи з 2000 р. було введене правило – забезпечувати описом Dublin Core практично кожну html-сторінку.

Модель електронної Інтернет-бібліотеки

Для моделювання інформаційних систем, зокрема функцій, які система повинна реалізувати, відношень між даними, поведінки системи в динаміці (аспектів реального часу) та структурного аналізу таких систем, серед множини засобів найбільш часто застосовуються два – **DFD** (Data Flow Diagrams) – діаграми потоків даних спільно зі словниками даних та їх специфікаціями, що є основним засобом моделювання функціональних вимог до системи, яка проектується, та **ERD** (Entity-Relationship Diagrams) – діаграми «сутність-зв'язок», призначені для розробки моделей даних, які надають стандартні засоби визначення даних і відношень між ними. Оскільки за допомогою ERD здійснюється деталізація сховищ даних проектованої системи, а також документуються сутності системи та засоби їхньої взаємодії, включаючи ідентифікацію об'єктів, важливих для предметної області (сутності), властивостей цих об'єктів (атрибутів) і їхніх відношень з іншими об'єктами (зв'язків), в роботі проектування БД, що відповідає за облік інформаційних ресурсів, виконано за допомогою ER-діаграми і подано на рис. 2.

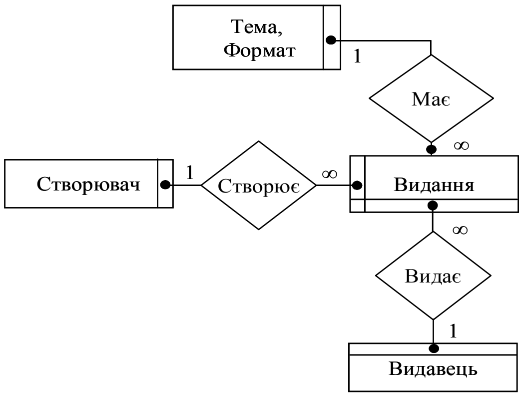


Рисунок 2 – ER-діаграма бази даних про інформаційні ресурси типу наукових видань

Відповідна даній ER-діаграмі схема бази даних наведена на рис. 3.



Рисунок 3 – Схема бази даних про інформаційні ресурси типу наукових видань

ER-діаграмі з 4-х сутностей відповідає база даних, яка у нормальній формі складається з 8 сутностей: видання, створювач, авторство, видавець, тип видання, мова, тема, формат.

Практична реалізація. Інтерфейс системи

Система «Віртуальна бібліотека» складається з двох частин: інтерфейсу наповнення та представлення на веб-сторінці у вигляді УДК та скрипта класифікації документа невідомої тематики за проіндексованими темами. Вважається, що користувачі, що матимуть доступ до наповнення бази документів УДК, будуть завантажувати в розділи тільки документи відповідної тематики. Завантажений об'єм документів буде еталонним для скрипта класифікації. Як бачимо (рис. 4), розділ «Віртуальна бібліотека» є однією з можливостей сайту Фізико-механічного інституту НАН України (http://www.ipm.lviv.ua/library/index.php).

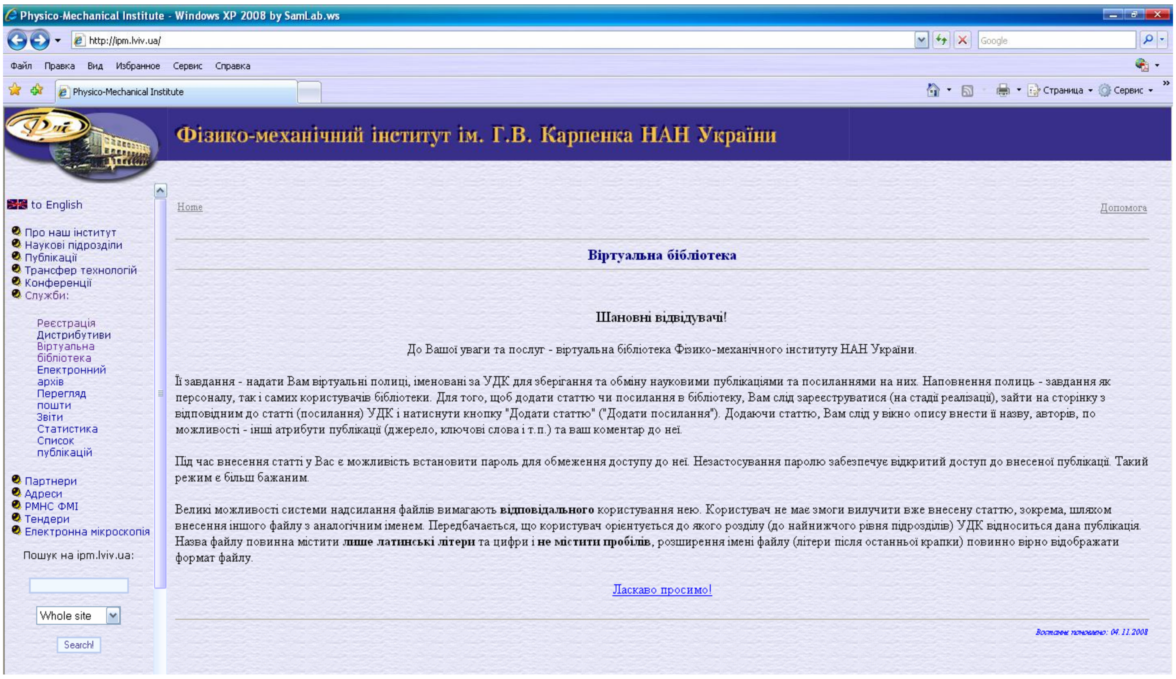


Рисунок 4 – Сторінка «Віртуальна бібліотека» на сайті ФМІ

Перша частина доступна через сайт, пункт меню «virtual library», при необхідності попередньо потрібно зайти у «registration» – реєстрація нових користувачів. Цей модуль через реєстрацію користувача дозволяє наповнювати той чи інший розділ, встановлювати права для документів (дозволити можливість «скачування» файлів тільки для зареєстрованих користувачів чи для всіх відвідувачів тощо), редагувати та видаляти за бажанням. Поділ на розділи проведено згідно з УДК (Універсальний Десятковий Класифікатор). Документи розміщені у файловій системі сервера інституту, інформація про систему розділів, користувачів зберігається у базі даних, бекенд якої використовує СКБД MySQL. Крім файлів документів інтерфейс також дозволяє вносити у розділи посилання на джерела інформації у глобальній мережі Internet. Є один користувач із абсолютними правами – адміністратор. Тільки він може додавати нові розділи та переглядати документи, завантажені іншими зареєстрованими користувачами незалежно від встановлених власником прав.

Для того щоб додати статтю чи посилання в бібліотеку, слід зареєструватися, зайти на сторінку з відповідним до статті (посилання) УДК і натиснути кнопку «Додати статтю» («Додати посилання»). Додаючи статтю, слід у вікно опису внести її назву, авторів, по можливості – інші атрибути публікації (джерело, ключові слова і т.п.) та коментар до неї (рис. 5). Під час внесення статті є можливість встановити

пароль для обмеження доступу до неї. Незастосування пароля забезпечує відкритий доступ до внесеної публікації. Такий режим є більш бажаним. Користувач не має змоги вилучити вже внесену статтю, зокрема шляхом внесення іншого файлу з аналогічним іменем. Це обмеження підвищує надійність функціонування підсистеми «Віртуальна бібліотека». Вжито усіх необхідних заходів для забезпечення сумісності з різними операційними системами та системами кодування в них. Інтерфейс служби простий та зрозумілий [8].

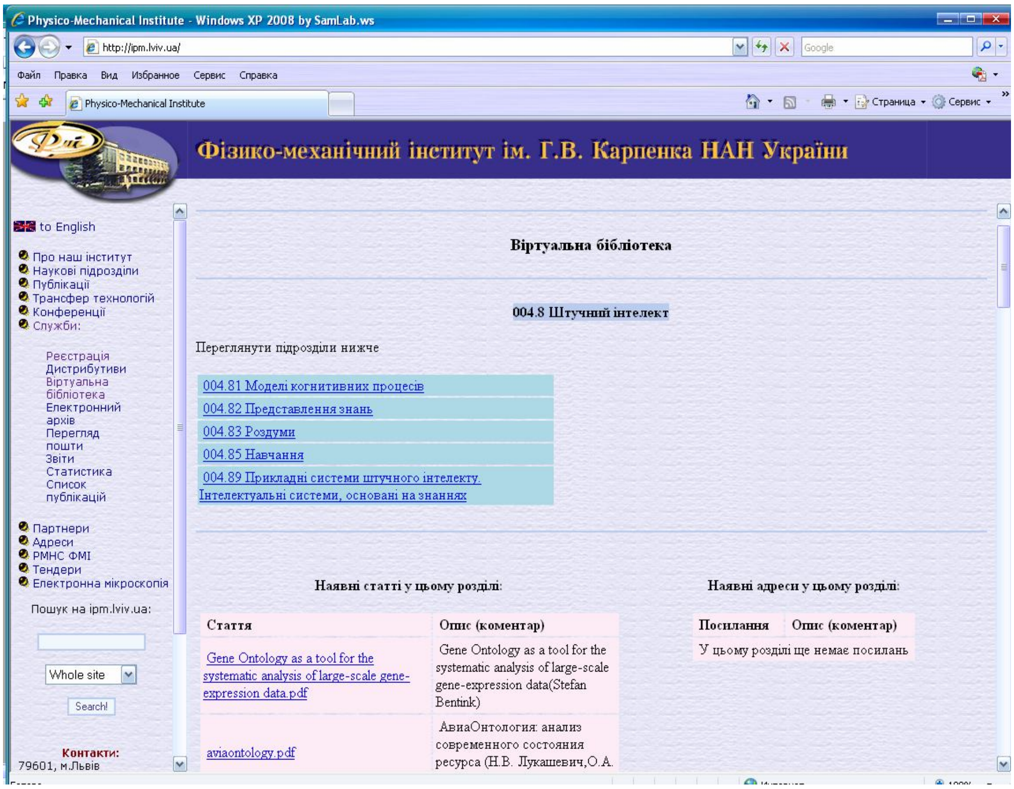


Рисунок 5 – Сторінка «Віртуальна бібліотека» на сайті ФМІ

Важливою функцією віртуальної бібліотеки як складової системи ВАРМ є забезпечення через глобальну комп’ютерну мережу зручного, оперативного, контекстно-залежного надання електронних документів та фотокопій першоджерел (електронна бібліотека).

Завдяки реалізації електронного каталогу віртуальної бібліотеки забезпечено виконання автоматизованої обробки нових надходжень в бібліотеку; звільнення співробітників бібліотеки від ряду рутинних робіт з підготовки картотек, видань, списків, замовлень, листів, звітної документації; створення бази даних про надходження. Автоматизація дозволяє виконувати пошук інформації по запитах читачів, обслуговування баз даних інформаційних і періодичних видань бібліотеки, введення масиву бібліографічних описів журнальних статей, пошук записів за ключовими словами, створення електронних довідників, контроль за виданою літературою, облік читачів і їх формулярів, автоматичний запис у формуляр читача виданої літератури, контроль терміну повернення книг. А головне – забезпечення читачам своєї бібліотеки віртуального каналу виходу у вітчизняний і світовий інформаційний простір. Бібліотеку ФМІ можуть наповнювати як персонал бібліотеки, так і користувачі. Для цього потрібно внести дані про видання (рис. 6) у відповідну форму інтерфейсу віртуальної бібліотеки.

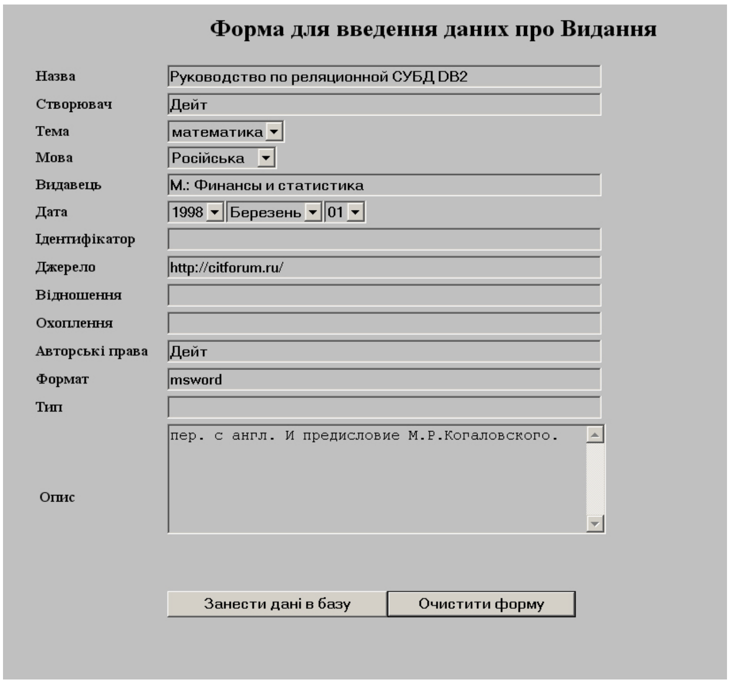


Рисунок 6 – Форма для введення даних про видання

Усі поля форми відповідають полям таблиці бази даних «Видання», причому обов’язковими є такі поля: Назва, Створювач, Тема, Мова, Видавець, Дата, Джерело, Формат. Доступ до цієї форми має лише бібліотекар. База даних «Видання» реалізована з використанням СУБД MySQL. Форма пошуку видань у віртуальній бібліотеці подана на рис. 7.

Тут користувач може здійснювати пошук за: а) назвою видання – для цього він може ввести повністю назву видання або ключові слова для пошуку; б) створювачем (автором) – це може бути один з авторів; в) темою; г) мовою; д) видавцем; е) датою.

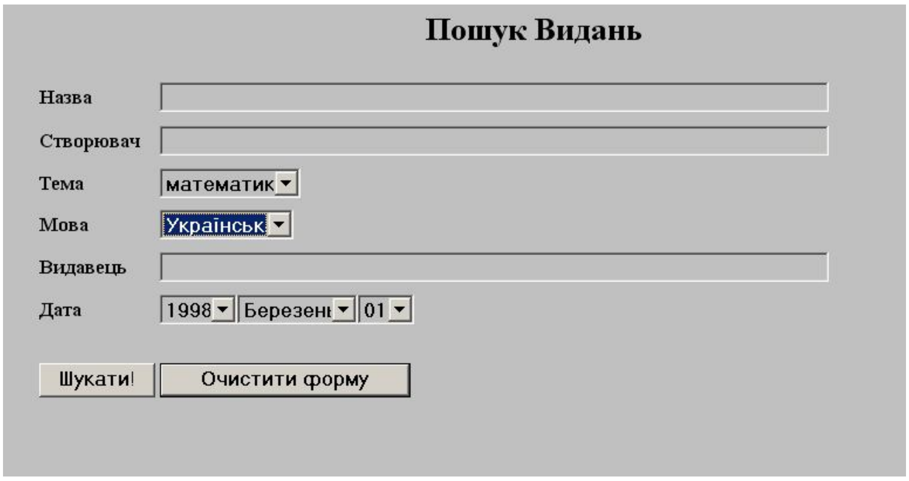


Рисунок 7 – Пошукова форма віртуальної бібліотеки

Основними полями для пошуку звичайно є перші два («назва» та «створювач»), решта чотири поля («тема», «мова», «видавець», «дата») є допоміжними. Допоміжні критерії дозволяють звузити параметри пошуку і таким чином зробити його швидшим та

ефективнішим, але вони не є самодостатніми, оскільки немає сенсу шукати всі книги, видані однією мовою. Для пошуку можна застосовувати як усі поля, так і деяку їх підмножину. Система дає можливість ведення ряду статистик: автори, що найбільше додають документів, нові надходження документів, найпопулярніші інформаційні ресурси.

Планується розробити модуль класифікації документів з інтегруванням його у загальний проект віртуальної бібліотеки. У ньому служба сама, використовуючи популярні пошукові машини, проводитиме попередній пошук за ключовими словами та «завантажувати» відповідно до заздалегідь вказаної квоти певну кількість найбільш відповідних до запиту публікацій, проводити класифікацію цих документів, визначати найбільш відповідні та видавати результати запиту користувачу віртуальної бібліотеки на його вимогу.

Висновки

Розроблена система може застосовуватись у будь-якій науковій установі для накопичення, впорядкування, зберігання та використання електронних документів.

Перевагами віртуальної бібліотеки є: швидкість пошуку потрібної інформації за кількома критеріями, пошук літератури за «ключовими словами», одночасний доступ до «популярного» видання, друкований примірник якого вже знаходиться в користуванні, доступ до раритетних екземплярів видань, доступ до яких у друкованому вигляді неможливий, збільшення кількості активних користувачів бібліотеки. Також спроектована система призначена для надання користувачам доступу до інформації про бібліотечні видання (монографії, довідники, збірників статей і т.п.), інформації про кожне видання, відомостей про книгообіг.

Те, що наукові статті та монографії будуть доступні в мережі Інтернет, допомагатиме студентам та аспірантам у навчанні, сприятиме професійному росту та визнанню самих науковців і, як наслідок, сприятиме розвитку вітчизняної науки.

Слід відзначити, що розроблена віртуальна бібліотека створює передумови для здійснення автоматичного навчання онтології метапошукової системи у відповідній до потреб користувача галузі (предметній області) в рамках ВАРМ наукового працівника.

Література

1. Вітенко В.І. Технологія створення електронної бібліотеки: підходи та перспективи [Електронний ресурс] / В.І. Вітенко // Всеукраїнська науково-практична конференція. – Режим доступу : http://library.kr.ua/conference/vitenko.html
2. Досин Д.Г. Розробка портальної експертної системи для автоматизованого пошуку науково-технічної інформації в мережі Інтернет / Д.Г. Досин, Р.Р. Даревич, М.М. Митроган // Відбір та обробка інформації. – 2003. – Вип. 18(94). – С. 123-126.
3. Даревич Р.Р. Віртуальне автоматизоване робоче місце з інтелектуальною компонентою / Р.Р. Даревич, Д.Г. Досин // Материалы 7-й ежегодной Междунар. конф. «Эффективность реализации научного, ресурсного и промышленного потенциала в современных условиях». – Киев : УИЦ «Наука. Техника. Технология». – 2007. – С. 69-71.
4. Даревич Р.Р. Створення експертної системи прогнозування пошкоджень в енергетиці та нафтохімії на основі металографічного аналізу / Р.Р. Даревич, А.Д. Марков, О.З. Студент // Машинознавство. – 2005. – № 2. – С. 48-53.

5. Клаус А.Ф. Совершенствование методов управления техническим уровнем и обновлением строительно-дорожных машин на основе создания АРМ инженера-исследователя : автореф. дис. … канд. техн. наук : 05.05.04 / А.Ф. Клаус. – М., 1990. – 25 с.
6. Пиотровская К.Р. Модели, программные и информационные средства учебного АРМ переводчика : автореф. дис. … канд.техн.наук : 05.25.05 / К.Р. Пиотровская. – К., 1993. – 17 с.
7. Лежебоков А.А. Автоматизированое рабочее место преподавателя с интеллектуальной поддержкой / А.А. Лежебоков, Л.А. Гладков // Программные продукты и системы. – 2005. – № 4. – С. 48-50.
8. Даревич Р.Р. Застосування інформаційних технологій для координації наукових досліджень / Р.Р. Даревич, Д.Г. Досин, В.В. Литвин. – Львів : СПОЛОМ, 2008. – 240 с.
9. Баркова О. Національна бібліотека України імені В.І. Вернадського [Електронний ресурс] / О. Баркова // Ресурси електронних бібліотек / Бібліотечний вісник. – 2001. – Вип. 1. – С. 20-24. – Режим доступу : http://www.nbuv.gov.ua/articles/2001/01bovreb.html
10. Dublin Core Metadata Initiative / DCMI Abstract Model. – Режим доступу: http://dublincore.org/documents/abstract-model
11. Обласна універсальна наукова бібліотека ім. Д.І. Чижевського [Електронний ресурс] / проект Мета дата. – Режим доступу : http://www.library.kr.ua/dc

Р.Р. Даревич, Д.Г. Досин, В.В. Литвин, А.А. Мазур

Виртуальная библиотека – элемент интеллектуальной системы для координации научных исследований

Статья посвящена созданию виртуальной библиотеки в связи с построением в Физико-механическом институте НАН Украины виртуального автоматизированного рабочего места научного работника. Электронный каталог базируется на стандарте метаданных Дублинское ядро (Dublin Core). Наполненная научными публикациями система служит основанием для осуществления автоматического обучения онтологии метапоисковой системы в соответствующей потребностям пользователя отрасли.

R. Darevych, D. Dosyn, V. Lytvyn, A. Mazur

Virtual Library as the Component of an Intellectual System for Coordinating of Scientific Investigations

The paper is devoted to construction of virtual library which elaborate the Karpenko Physics-Mechanical institute of NAS of Ukraine of virtual researcher workplace. The electronic catalog is based on the Dublin Core metadata standard. The library filled with scientific publications creates pre-conditions for realization machine learning of the meta-search system ontology in the user domain within the framework of the researcher virtual workplace.

Стаття надійшла до редакції 20.07.2009.